Fugitive Information on the World Wide Web: A cost Effective Method of Access for a Diverse Clientele 1

Lisa G. Dunn
Arthur Lakes Library, Colorado School of Mines

Joanne V. Lerud
Arthur Lakes Library, Colorado School of Mines

FUGITIVE INFORMATION ON THE WORLD WIDE WEB:

A COST-EFFECTIVE METHOD OF ACCESS FOR A DIVERSE CLIENTELE

Lerud, Joanne V., and *Dunn, Lisa G*.
Arthur Lakes Library, Colorado School of Mines,
Golden 80401, Colorado, USA

Introduction

Libraries rely principally on traditional methods of bibliographic control, such as cataloging and indexing, to provide access to their core collections. However, even the most focused library has collections of materials for which traditional methods of bibliographic control are not appropriate. These collections often are peripheral to the central mission of the library or would be too resource-intensive to catalog. For various reasons, though, discarding them is not possible. These collections become "fugitive information" within the library, available only to those patrons willing to make the effort to find them. Although libraries have not generally received increased funding over the last few years, new technological applications have made it possible for us to look at these collections in a new light. While we cannot necessarily do more with less, we can do what we do differently.

Overview of Fugitive Information

The term "fugitive information" refers to information that is difficult to identify despite informed use of standard indexes and bibliographic databases, or is difficult to acquire through standard publishing channels. Whether one considers information "fugitive" depends in part on local circumstances: the indexes available to the user; the scope of information covered by those indexes; availability of materials locally or through interlibrary loan, and so forth. Almost all marginally available information can be fugitive, but some collections of information are more fugitive than others.

Subject collections containing a high percentage of gray literature fall within the definition of fugitive information. Discussions of gray literature and the corresponding problems of identification and acquisition appear in both the library literature and that of other disciplines (Wood 1984; Bichteler 1991; Debauche 1995; Meier and Greenberg 1994). Gray literature is characterized by limited distribution or distribution outside conventional publishing mechanisms, poor bibliographic control, and non-standard formats. It is generated quickly, distributed as an in-depth response to specialized information needs, and directed to a specific, frequently small audience.

Other subject collections provide different acquisition and access challenges than

those common to gray literature. Highly specialized, interdisciplinary, or multidisciplinary materials, even if published through standard channels or included in common indexes and catalogs, may not be easily identified as belonging to a particular subject area. Subject-specific indexes such as Chemical Abstracts or Engineering Index provide extensive controlled vocabularies for expert keyword coverage of highly specialized materials, but may not adequately represent multidisciplinary literature, and rarely cover gray literature. The equivalent electronic databases, although permitting keyword searching on author-generated titles and abstracts, mirror their print versions in the exclusion of out-of-scope materials. Library public access catalogs (PACs), while theoretically unlimited in subject scope, are drastically limited by both a lack of abstracts or tables-of-contents and a more general controlled vocabulary.

To locate fugitive information, the user typically must consult many different indexes and catalogs (in addition to information professionals), making broad generalizations for complete coverage of the subject, which result in high retrieval but low relevance in search sets. Once materials are identified, the user must gather them through the peer network, direct contact with authors or publishers, and multiple interlibrary loans. Obtaining fugitive information is not impossible, but use of formal channels can be time consuming and costly, one reason why the "invisible college" of peer networks is so much more popular with patrons than simply using the library's services (Bichteler and Ward 1989). Users' difficulties in obtaining fugitive information often drive librarians to create small subject collections for their local users.

Local subject collections help solve local users' problems, but even these collections bother librarians because of inadequate access and disproportionate drain on resources. The collections are needed but not used enough to justify the expenditures that would make them more accessible. Since information about a collection isn't disseminated, users stumble on it via the peer network or chance contact with the librarian. The patron population for which such a collection has been established may be localized either geographically or by discipline, but the information may be extremely valuable to others. However, potential users from remote geographic regions or other disciplines are even less likelt to know of its existence. Even if they do discover its existence, access (physically or bibliographically) may not be possible. Lack of widespread demand for these collections, disparity in formats (monograph vs. serial article, for example), and the lack of resource commitment necessary to adequately catalog or index them means that as "collections" they remain on the fringes of the information resource world.

The Arthur Lakes Library of the Colorado School of Mines is one of the many libraries that has found it necessary to maintain local collections of fugitive information for its patrons. Some of these collections are composed primarily of gray literature; others contain more conventional materials, but as collections are unique to this Library. All of these collections have in common a specialized subject focus and all present challenges to bibliographic control.

Arthur Lakes Library and Its Collections

The Colorado School of Mines (CSM) is a public research university with broad

expertise in natural resource exploration, extraction, production, and utilization. The School offers the Bachelor of Science degree in engineering, mathematics, and the physical sciences. At the master's and doctoral levels, degrees are offered in mineral economics, engineering, mathematics, and the physical sciences. In addition, the School maintains active research programs in academic departments and affiliated research institutes.

The Arthur Lakes Library supports the educational and research programs of the Colorado School of Mines as well as the needs of an external community of users at the local, regional, and international levels. The Library makes its collections of monographs and serials accessible via CARL PAC, the Library's public access online catalog, and OCLC, an international bibliographic utility. Collections that have not been traditionally (or thoroughly) cataloged on CARL and OCLC include mine reports, company annual reports, the Tell Ertl oil shale collection, historical photographs, Society of Mining Engineers preprints, and wire rope materials.

Mine Reports

Mine reports are quasi-formal publications generated by state inspectors of mines, companies, consultants, or individual mining engineers. They can include maps, assay information, site particulars, descriptions of ore and host rock, and descriptions of day-to-day site operations. They are invaluable both for their information on past and present mining activities and for their economic and environmental data. Not surprisingly, many mine reports describing Colorado mines have found their way to the Arthur Lakes Library by gift or deposit, and others continue to do so. Historically in our Library, mine reports could be found in the cataloged/classified monograph collection, in archival manuscripts collections, in the map collection, and as a separate collection of their own.

Persons interested in this type of information might guess that it could be found at the Colorado School of Mines, and might even find a few scattered reports in our PAC, although few reports were initially cataloged. Since these reports are unique, catalog records for them rarely appear in bibliographic utilities. Inquiries concerning the mine reports had to be directed to a reference librarian, who would refer the patron to either a bound index or a card file index. If a reference librarian was not available, or if the patron had no prior knowledge of the means of access, the information would not be found. Because of preservation concerns, the archival collection of mine reports was eventually copied to microfilm and the paper reports closed to the public. This investment in the collection is an indication of its continuing value. However, it introduced another twist to the problem of access, since the microfilmed reports were identified only by roll and frame numbers, which in turn were indicated only by obscure markings on the card file index.

Company Annual Reports

Company annual reports and supplementary materials such as 10-Ks have long formed a part of the Library's holdings. This collection focuses on national and international minerals resource companies. The Library currently houses about 1,500 titles, some dating back to the early 1900's. These reports serve as primary information sources for the mineral economics curriculum and for mineral, environmental, and energy-related industries. CSM students also review company literature prior to job interviews. Acquisition is by gift and by direct request from the

company. Access to this collection has been provided through a simple card file showing company name and holdings. At times when resources were particularly short, acquisition and maintenance were spotty.

In addition to these two collections, the Library maintains several others with even more limited bibliographic control. Because of their characteristics and the Library's finite cataloging resources, it was decided to provide bibliographic access through separate database management software called Personal Librarian (PL). PL was originally a text-based program that performed natural language searching and ranked results according to a relevance algorithm. PL allowed us to inventory and index these collections at whatever level we chose. The PL databases made the collections accessible locally, to library staff and patrons who inquired, and to anyone who had the technical capabilities to connect to PL by a telnet session through the Internet. Access, therefore, was not an intuitively obvious process, but it was better than that provided by a local card file.

Tell Ertl Oil Shale Repository

The Tell Ertl Oil Shale Repository, 19 discrete collections of materials, consists of published and unpublished manuscripts, photographs, slides, and maps on the development of the oil shale industry throughout the world, with emphasis on the western United States. Materials range from general documentation to highly technical works. The Repository materials are managed as an archival collection of historical interest chronicling oil shale's four boom and bust cycles in the American West. It is a dead collection, with no plans for further acquisitions, so it would be difficult to justify allocating full cataloging and processing resources for access and retrieval.

In 1990, the Tell Ertl Family Trust provided funding for supplies, room appointments, student assistance, and limited remuneration for a subject specialist. His efforts were used to organize, inventory, and index this collection using PL.

Colorado School of Mines Historical Photographs

The Photographs collection is an eclectic mix of photographs, slides, negatives, stereopticons, and lantern slides dating back to the 1880s and focusing on people, buildings, and activities associated with the School of Mines. Also included are historical mining scenes, regional scenic panoramas, and classroom teaching aids. Over time, few pains were taken with preservation, and the collection was housed unindexed and uninventoried in filing cabinets, boxes, and envelopes. The Library received a small but steady stream of requests to use these materials, but scholars were forced to review the entire collection while looking for photographs of interest.

The Photographs collection was considered a good candidate for a PL database, and as a result, minimal descriptive information for each item was entered into PL beginning in 1991. Preliminary preservation measures such as acid-free envelopes were also undertaken at this time.

Society of Mining Engineers Preprints

Each year, the Library receives preprints from the Society of Mining Engineers. This collection dates back to 1959 and is stored in filing cabinets by preprint number. Print indexes of SME publications are available, but are not always satisfactory; they may

not be current or inclusive, or they may not offer the keyword searching that is sometimes necessary for these technical subjects. Starting in 1993, Personal Librarian was used to develop a bibliographic index of SME preprints, including back issues.

Information Center for Ropeway Studies

In 1991, the Information Center for Ropeway Studies was established at the Arthur Lakes Library. The Center's collection contains information on the history, theory, design, and operation of ropeway systems and wire rope. It began with the gift of the working library of Charles Dwyer, a Colorado ropeway engineer, and now includes approximately 250 monographs as well as journals, newsletters, manufacturers' catalogs, reprints, videos, slides, photographs, and lantern slides. The monographs and journals are cataloged and accessible via CARL. Journal tables of contents from 1988-present are included in UnCover, a journal index/document delivery database also available through CARL. All other materials are categorized as gray literature. The Center's collection is being added to actively, funded in part by an external organization, OITAF-NACS (International Organization for Transportation by Rope - North American Continental Section), not affiliated with the Colorado School of Mines.

From its inception, the Ropeway Center's Advisory Board desired to have ropeway information available both nationally and internationally, and at greater depth than that provided by a general online catalog or table-of-contents database. In addition, although the Center was housed at CSM and had local scholarly users, the Board assumed that the large majority of users of the Center would not be on site. What the Board really wanted was a publicly accessible, comprehensive database of all materials in the Center and elsewhere. PL was selected as the means by which to address this need.

Impact of the Internet on Bibliographic Control Decision-Making

The expectations of our students, faculty, and external users have necessitated a review of the fugitive information held within the Library for the purpose of easier, faster, and wider access, and access to the desktop. In 1991, the Library began an overall assessment of its collections to determine the appropriate level of resource commitment. This ongoing assessment is based on library mission, patron population, and collection development policies. The Library's fugitive collections came under particular scrutiny because of the problems inherent in maintaining and providing access to them. Advances in technology played a major role in the resolution of those deliberations.

Our public access catalog on CARL first became accessible by Internet telnet session in 1992, with CARLWeb, the software's World Wide Web (WWW) interface, appearing in 1996, greatly improving our PAC's visibility and accessibility. While our cataloged materials have been accessible through OCLC since 1975, until recently this bibliographic utility had been used primarily for technical functions by library staff. The availability via the Internet of OCLC's FirstSearch WorldCat database, an end-user gateway to OCLC's bibliographic database, changed this. Remote end-users now have WWW access to the Arthur Lakes Library's cataloged materials not only through CARL but through WorldCat as well. Everything that we catalog now

reaches a wider audience.

The developers of PL, our text-based database management software, have released a web-based user interface, PLWeb. Our library purchased the PLWeb software in 1996. This product's search engine is very flexible, allowing both the novice and the expert to search for information effectively. All fields are searchable and search results appear as a ranked list of documents. PLWeb's advanced search features include:

- Stem searching: the term is stripped of common suffixes and the search is performed for variants of the resulting root word
- Concept searching: searching based on a general concept
- Feedback searching: results from one search can be incorporated into the next search
- Relevance ranking: ordering of search results by the frequency of distinct search terms in a retrieved document and by frequency of each of those terms in the whole database

Our primary patron population, CSM users, desires information delivered to their desktops and the Internet can provide this. The Internet also provides the potential to increase our external patron population. These two tools (CARLWeb and PLWeb) are well-suited to our Library's operations and have required us to re-assess our efforts to make available our collections of fugitive information. In some cases, we diverted resources to handle materials traditionally by cataloging. In other cases, Personal Librarian was selected as the more appropriate medium for bibliographic access.

Criteria for Access

Should it be done? The first question that needs to be asked is, *"Should bibliographic access to this collection be available through the Internet?"* Despite the urge to make true the fantasy that everything is on the Internet, the answer to this question is not always yes. Information may already be on the Internet or be accessible through a common printed publication. We determined for the purpose of collection development that, since our resources are limited, our library's contributions to Internet-accessible information should be unique and not likely to be duplicated by others. Our collections of fugitive information make good candidates for Internet access because of their uncommon or unique nature.

The repercussions of Internet access need to be considered as well, since there is little point in providing bibliographic access if the remote user cannot actually use the materials. As part of our collection assessment, our Document Delivery section has been developing processes over the past five years to improve turn-around time and accommodate requests for special materials. We can now fill most requests in a timely manner and at a reasonable price. Media such as videos and photographs still present challenges, however.

Should we do it? The second question to ask is, *"If this information should be on the Internet, should WE be the ones to do it?"* At this point, mission statement and capabilities enter into the picture. As an academic institution, our primary mission is to support the educational and research efforts of the Colorado School of Mines. We also assist external users, but to divert significant resources to these users without

external funding would be inappropriate. Because of our technological infrastructure (support from the campus computing center), we have decided to use the Internet as the best means of improving access to our fugitive collections for our primary user population, which has expressed the need for improved off-site access to our Library's information. That this information is available to the world is an added bonus.

Can we do it? The bottom line for making fugitive information accessible through the Internet is not only *"Should we do it?,"* but *"Do we have the resources to do it?"* Technological infrastructure and bibliographic utility tools are vital, but as every librarian knows, you can have these means and still not be able to provide access due to lack of staff or funding. We have to determine whether it is worth the investment to provide access to fugitive collections via the Internet. Collection scope, potential user population, and availability of staff expertise must be considered. Even with the technological resources available, a resource commitment will still impact a library's operations. We determined that we would only provide Internet access for those materials which would be a permanent part of our collections, those carrying archival value.

The actual method chosen for bibliographic control depended on a variety of factors:

- How big was the collection?
- Was the collection growing or static, and if growing, how fast?
- Are we planning to expend additional resources for necessary preservation or maintenance?
- Do the collection formats lend themselves better to full-level MARC record cataloging or to
  Personal Librarian protocols?
- Would original cataloging be necessary?
- Is there external funding or support available for this work?
- Do we plan on additional access features (graphics, full text, etc.)?

Our fugitive collections were determined worth the effort, some more so than others. Provision of bibliographic control, whether through MARC-record library catalogs or custom-built databases, is a value-added process: subject headings, abstracts, notes, and document delivery information can be added in either situation. However, traditional practices, costs, and user expectations must be considered in selecting the appropriate value-added process.

Cataloging on CARL/OCLC

As a School of Mines with an international reputation and a unique and valuable collection of mine reports, we felt that this information should be made accessible. We made the decision to catalog all of our mine reports in the public access catalog. This collection grows slowly, and our cataloging expertise is appropriate for the subject matter. Since the information has historical value, there are no plans to discard this material. Current personnel, temporary additional personnel, and associated OCLC cataloging charges were deemed appropriate investments to make this collection of gray literature accessible to any interested person.

The company annual reports collection was re-assessed in 1996; the process upheld the importance of this collection as a primary information resource for the curriculum.

Again, the Library's clientele expect that such a collection would be on site. This collection, unlike the mine reports, is a rapidly growing collection, with both new items and new titles. It is considered regionally unique. Preliminary surveys of the OCLC database indicated that these companies' reports are rarely formally cataloged (Stark 1997). The collection contains several titles over 50 years old, and we plan to retain all materials and make the collection an archival record of minerals company activity. Resources were expended for appropriate maintenance, staffing for cataloging and research, and OCLC charges. We will continue to catalog this collection as long as our resources permit.

Indexing on Personal Librarian (PLWeb)
The Tell Ertl Collection, Photographs Collection, SME Preprints, and Ropeway Center Collection were already accessible via Personal Librarian. The decision to switch to PLWeb was relatively straightforward from a resource standpoint.

Tell Ertl is an example of a collection that may indeed remain at a minimal level of electronic access. It is a dead collection with a large number of discrete items. Most items would require original cataloging; many are in non-standard formats. The oil shale industry is inactive at present, and collection use is low. Should the oil shale industry go through another "boom" cycle, access practices will have to be re-assessed. Each PL Tell Ertl record contains: collection name; accession number; descriptors; title; document type; primary source; document length; date; author(s). The collection and the Library benefit from this low-cost way to make this information available to its user population.

The Photographs collection is now completely entered into PL. This collection still receives limited but steady use. With WWW access, potential users include historians in western US history and the history of US higher education, CSM alumni, and genealogists. PL records contain: collection number; photographer; title; year; medium; color; size; notes; donor fields. If resources become available, this collection is a prime candidate for the mounting of digital files for each photo on the WWW. However, until this is accomplished, the basic descriptive information is accessible.

The SME preprints' records are added to PL each year as new material is received. The preprints have a local user population of mining and civil engineers, metallurgists and chemists. WWW access will improve the preprints' availability to members of these professions elsewhere. PL records include: preprint number; title; author(s); date. The basic bibliographic information provided in the current database is adequate for most queries, but more searchable text would improve retrieval rates. We are in the process of adding abstract fields to the current year's papers, and plan to add abstracts retrospectively as resources permit.

The Ropeway Center Advisory Board's desire for a comprehensive, remotely accessible database is driving development of its PL database. The Ropeway database now includes not only those materials on site at the Center, but information housed elsewhere on ropeway systems. Records include: author; title; source; meeting; publisher; publication year; keywords; abstracts; location; call number; language. ROPEWAY is a value-added database since volunteer subject specialists have supplied abstracts and keywords for these fields to enhance retrieval by using both natural language and controlled vocabulary.

Assessment

The formal cataloging of mine reports and annual reports requires an investment of additional staff time and operations dollars. Library users now retrieve mine reports and annual reports as a regular part of queries to CARL, which is generally more efficient for patrons, on site or remote, as well as for reference staff. Physical control of these items is the same as that of our working collection, with security tape, labeling, and shelf maintenance.

The annual report collection now receives a new level of management. In the past, spotty attention to the mailing list resulted in gaps. This collection will now be provided with the same acquisition control as other serial publications, including electronically generated claims. Processing procedures will also be the same as for the working collection. Qualified staff will ensure appropriate processing using occasional student help but not making the students responsible for professional level decisions..

Librarians know that making information adequately (much less fully) accessible is not a simple process. We, at the Arthur Lakes Library, are still learning what it takes to accomplish our objectives with our fugitive information and living with our decisions. Expanding the scope of formally cataloged collections has obvious and immediate impacts on technical services staff. We helped address this problem by using temporary staff and by cross-training permanent staff whose current responsibilities and aptitudes made this feasible. The Library's flat organizational structure made this process relatively easy.

Particular strengths of PL include a powerful search engine and uncomplicated database construction. We have in the past tailored record entry to allow student assistants and volunteers to do the bulk of the data entry for some collections. However, increasing the number of participants building the files also increases both the difficulty of communicating rules for standardization and the attendant amount of internal documentation. Lack of control in this regard resulted in a significant amount of unnecessary database cleanup. We are in the process of correcting this problem, developing simple but consistent standards for data entry, and writing and distributing documentation. PLWeb allows us to track hits for each database; we will be able to assess site use and compare that to use of the collections.

Conclusion

Our fugitive collections have a small user population amongst our primary users. The subjects are within the scope of our collection development areas, although highly specialized within those areas. They have been in effect underused, even though unique. The Internet and more specifically the World Wide Web have characteristics that suit us very well in the areas of bibliographic control and access. The ability to make our collections accessible through the Internet is shifting our priorities, primarily in the area of service to our patron population. Given the comparatively large population of Internet users, our potential user population for these collections grows from small and local to large and scattered. By *"scattered,"* we mean not just geographically, but by subject discipline. (As the Internet's search engines develop,

this information should become increasingly available to those from other disciplines searching for multidisciplinary or interdisciplinary information.). Both user populations are important to us. Although altruism does play a part, we expect benefits from this process. Improvements in access through the Internet act as a marketing tool for these collections. By generating more remote use, we improve the possibility for more advantageous reciprocal agreements for our library and our primary user population. We give the Library and the campus a higher profile to the public and to the research community. This in turn creates the potential for increased funding and gifts of specialized materials to the collections, improving their coverage and, again, our primary user population's access and resources.

Acknowledgments:
The authors are grateful to Janice Christopher for her editing efforts.

References Cited:

WOOD, D. N. The collection, bibliographic control and accessibility of grey literature. IFLA Journal 10(3) 1984: pp.278-282.

BICHTELER, Julie. Geologists and gray literature: Access, use, and problems. Science and Technology Libraries 11(3) 1991: pp.39-50.

DEBACHERE, M. C. Problems in obtaining grey literature. IFLA Journal 21(2) 1995: pp.94-98.

MEIER, Alan and GREENBERG, Steve. A journey through the gray literature. Home Energy 11(4) 1994: pp.17-21.

BICHTELER, Julie and WARD, Dedrick.. Information-seeking behavior of geoscientists. Special Libraries 80(3) 1989: pp.169-178.

STARK, Marilyn. Personal communication. Golden CO: Arthur Lakes Library, Colorado School of Mines. May, 1997.

Acronyms and Abbreviations:

CARL	Colorado Alliance of Research Libraries (integrated library system)
CARLWeb	Colorado Alliance of Research Libraries web interface
CSM	Colorado School of Mines
MARC	Machine Readable Cataloging
OCLC	Online Computer Library Center (bibliographic utility)
OITAF-NACS	International Organization for Transportation by Rope-North American Continental
PAC	Public Access Catalog
PL	Personal Librarian

PLWeb	Personal Librarian web interface
SME	Society of Mining Engineers
US	United States
WWW	World Wide Web

Biographical Information

Joanne Lerud is the Director of the Library, Arthur Lakes Library, at the Colorado School of Mines. She received her B. S. and M. S. in Geology from the University of North Dakota and her M. A. in Librarianship and Information Management from the University of Denver. She has worked in industry as well as academic library administration.

Lisa Dunn is Head of Reference, Arthur Lakes Library, at the Colorado School of Mines. She received her B. S. in Geology from the University of Wisconsin-Superior, her M. A. in Geology from Washington University, and her M. L. S. from Indiana University. She has held reference and administrative positions at several academic posts.

Both authors can be reached at the Arthur Lakes Library, Colorado School of Mines, Golden, Colorado 80401. They may be reached via e-mail at jlerud@mines.edu or ldunn@mines.edu.

The Library's web page can be reached at: http://www.mines.edu/library/